On the Aboutness of UNL

Ronaldo Teixeira Martins,1,2 Maria das Graças Volpe Nunes.2,3

1Faculdade de Filosofia, Letras e Educação – Universidade Presbiteriana Mackenzie
Rua da Consolação, 930 – 01302-907 – São Paulo – SP – Brazil

2Núcleo Interinstitucional de Linguística Computacional (NILC)
Av. Trabalhador São-Carlense, 400 – 13560-970 – São Carlos – SP – Brazil

3Instituto de Ciências Matemáticas e da Computação – Universidade de São Paulo
Av. Trabalhador São-Carlense, 400 – 13560-970 – São Carlos – SP – Brazil

ronaldomartins@mackenzie.com.br, gracan@icmc.usp.br

Abstract. This paper addresses the current status, the structure and role of the
UNL Knowledge Base (UNLKB) in the UNL System. It is claimed that the
UNLKB, understood as the repository where Universal Words (UWs) are
named and defined, demands a thorough revision, in order to accomplish the
self-consistency requirement of the Universal Networking Language (UNL). In
order to emulate human cognition and constitute the “aboutness” of the UNL,
the UNLKB should be decentralized, distributed and reorganized as a network
of networks, allowing for multicultural information and dynamic data.

1 Introduction

The Universal Networking Language (UNL) is an “electronic language for computers
to express and exchange every kind of information” [Uchida, Zhu & Della Senta,
1999]. It can be defined as a knowledge-representation formalism expected to figure
either as a pivot language in multilingual machine translation (MT) systems or as a
representation scheme in information retrieval (IR) applications. It has been devel-
oped since 1996, first by the Institute of Advanced Studies of the United Nations Uni-
versity, in Tokyo, Japan, and more recently by the UNDL Foundation, in Geneva,
Switzerland, along with a large community of researchers—the so-called UNL Soci-
ety—representing more than 15 different languages all over the world.

Formally, the UNL is a semantic network believed to be logically precise, hu-
manly readable and computationally tractable. In the UNL approach, information
conveyed by natural language utterances is represented, sentence by sentence, as a
hyper-graph composed of a set of directed binary labeled links (referred to as “rela-
tions”) between nodes or hyper-nodes (the “Universal Words”, or simply “UW”),
which stand for concepts. UWs can also be annotated with attributes representing con-
text-dependent information.

As a matter of example, the English sentence ‘Peter kissed Mary?!’ could be rep-
resented in UNL as (1) below:

© J. Cardeñosa, A. Gelbukh, E. Tovar (Eds.)